Controllability Metrics on Networks with Linear Decision Process-type Interactions and Multiplicative Noise

Tidiane Diallo, Dan Goreac

Controllability Metrics on Networks with Linear Decision Process-type Interactions and Multiplicative Noise

Tidiane Diallo* Dan Goreac*†

October 12, 2015

Abstract

This paper aims at the study of controllability properties and induced controllability metrics on complex networks governed by a class of (discrete time) linear decision processes with multiplicative noise. The dynamics are given by a couple consisting of a Markov trend and a linear decision process for which both the "deterministic" and the noise components rely on trend-dependent matrices. We discuss approximate, approximate null and exact null-controllability. Several examples are given to illustrate the links between these concepts and to compare our results with their continuous-time counterpart (given in [16]). We introduce a class of backward stochastic Riccati difference schemes (BSRDS) and study their solvability for particular frameworks. These BSRDS allow one to introduce Gramian-like controllability metrics. As application of these metrics, we propose a minimal intervention-targeted reduction in the study of gene networks.

AMS Classification: 93B05, 60J05, 90C40, 93E03, 92C42

Keywords: linear decision process; null-controllability; backward stochastic Riccati difference scheme; controllability metric; gene networks; phage λ

1 Introduction

We focus on a particular class of discrete-time decision processes described by a couple denoted (L, X) and consisting of a Markovian trend and a linearly trend-based updated component. This kind of processes naturally appear in the study of complex systems (such as regulatory gene networks). In this setting, the trend component corresponds to a finite family of DNA configurations which induce regime changes on functional components (usual proteins) X. Decisions are assumed to be made at expression level in order to obtain suitable behavior of X. We try to give a mathematical answer to the following questions. Given a family of possible actions, what are the minimal interventions to be selected in order to guarantee a targeted response. Second, how can this be quantified through a metric at the level of biochemical reactions network ? To answer these questions, we envisage invariance and Gramian-type descriptions of controllability concepts. This paper can be seen as a discrete-time counterpart of [16] in which piecewise deterministic Markov processes of switch type are considered. Together, the papers cover the two usual points of view over controlled switch processes with linear updating : the averaged, piecewise deterministic (macroscopic) perspective (in [16]) and the marked point process (closer to microscopic perspective in this paper).

The process L is assumed to be a finite-state Markov process on a filtered probability space $(\Omega, \mathbb{F}, \mathbb{P})$, with transition measure Q and taking its values in $\mathcal{B} = \{e_1, e_2, ..., e_p\}$, for some integer

*Université Paris-Est, LAMA (UMR 8050), UPEMLV, UPEC, CNRS, F-77454, Marne-la-Vallée, France, Dan.Goreac@u-pem.fr

†**Acknowledgement.** The work of the second author has been partially supported by he French National Research Agency project PIECE, number **ANR-12-JS01-0006.**

$p \geq 2$. Without loss of generality, Ω is set to be the discrete sample space $\Omega = \mathcal{B}^{\mathbb{N}}$ and we assume the filtration $\mathbb{F} = (\mathcal{F}_n)_{n\geq 0}$ to be the natural one associated to L. Following [5], [3] and without loss of generality, we assume that $\mathcal{B}$ is the standard basis of vectors of $\mathbb{R}^p$. We introduce the martingale M by setting

$$M_n := \sum_{k=0}^{n-1} \left[L_{k+1} - \mathbb{E}\left[L_{k+1} / \mathcal{F}_k \right] \right],$$

for $1 \leq n \leq N$, with the obvious convention $M_0 = 0$. As usual, we let $\Delta M_n := M_n - M_{n-1}$, for all $n \geq 1$. We will be focusing on a class of decision processes X on some state space $\mathbb{R}^m$, for $m \geq 1$ and controlled by d-dimensional processes, for some $d \geq 1$. The evolution is given by linear updating and multiplicative noise

$$(1) \qquad \begin{cases} X_{n+1}^{x,u}(\omega) = A_n(\omega) X_n^{x,u}(\omega) + Bu_{n+1}(\omega) + \sum_{i=1}^{p} \langle \Delta M_{n+1}(\omega), e_i \rangle C_{i,n}(\omega) X_n^{x,u}(\omega), \\ X_0^{x,u} = x \in \mathbb{R}^m. \end{cases}$$

The process A is $\mathbb{R}^{m\times m}$-valued and $\mathbb{F}-$adapted, the matrix $B \in \mathbb{R}^{m\times d}$, and C_i are $\mathbb{R}^{m\times m}$-valued and $\mathbb{F}-$adapted processes for all $1 \leq i \leq p$. The $\mathbb{R}^d$-valued control process u is taken to be square-integrable $\mathbb{F}$-predictable. The set of all $\mathbb{F}$-predictable processes is denoted by $\mathcal{P}red$. Whenever no confusion is at risk, we will drop the dependency on ω. The reader may want to note that this provides a slightly more general framework than Markov decision processes since the coefficients are adapted (i.e. functions of the time parameter n and the vector $(L_0, L_1, ...L_n)$). On the other hand, the transition measure has a particular form.

The first aim of the paper is to characterize controllability properties for systems driven by (1) i.e. the possibility to direct the process towards a coherent target. For controlled linear deterministic systems $\dot{X}_t = AX_t + Bu_t$, the controllability properties are summarized by the celebrated Kalman criterion stating that $Rank\left[B\ AB\ A^2B\ ...A^{m-1}B\right] = m$. Similar assertions are valid for discrete systems $X_{n+1} = AX_n + Bu_{n+1}$. This can equally be extended for Markov decision processes driven by non-random coefficients and additive noise of type $X_{n+1} = AX_n + Bu_{n+1} + \xi_{n+1}$. However, for continuous-time controlled linear systems with multiplicative stochastic perturbations, this condition is no longer sufficient. For examples pointing to this direction, the reader is referred to [2] or [14] (for Brownian perturbations), [15] or [16] (for continuous-time switch processes).

One can, alternatively, study the dual notion of observability via Hautus' test as in [18]. The criteria involve algebraic invariance notions which are independent of the space on which they are studied. For infinite-dimensional settings, the reader is referred to [28], [8], [27], [20], [19], etc.

In the continuous-time stochastic setting, duality techniques lead to backward stochastic differential equations (BSDE introduced in [24]). With these tools, exact (terminal-) controllability of Brownian-driven control systems is linked to a full rank condition in [25]. When the control is absent from the noise term, one studies approximate controllability, resp. approximate null-controllability. Invariance criteria are given in [2] for the control-free noise and [14] for the general Brownian setting. In the case when the stochastic perturbation is of jump-type, exact controllability of continuous-time processes cannot be achieved. This follows from incompleteness (cf. [22]) and one has to concentrate on approximate controllability.

For continuous-time control systems with Brownian noise, approximate and approximate null-controllability notions coincide (cf. [14]). This is no longer the case (see [12]) when an infinite-dimensional component of mean-field type governs the Brownian-driven systems. Various methods can be employed in infinite-dimensional state space Brownian setting leading to partial results (see [9], [29], [1], [11]).

The main goal of the first part of our paper is to study the controllability properties of the Markov decision process with linear updating and multiplicative noise perturbations. It can be seen as a discrete-time counter-part of [16] and, to some extent, [29]. We begin with a duality result between controllability and observability in Section 2.1. This leads to considering some adjoint

process satisfying a backward difference scheme. Its construction is close to backward stochastic difference equations (see, for example, [5], [3], [4]). The first main result of the paper (Theorem 2) gives two characterizations for approximate null-controllability and a duality criterion for approximate controllability. It equally states the equivalence between approximate and exact null-controllability. However, unlike the continuous-time frameworks (compare with [2] for Brownian systems and [16, Section 4.1, Criterion 3] for jump-systems), in discrete-time, null-controllability does not imply approximate controllability. This surprising behavior is illustrated in Example 5.

To construct a controllability metric, we concentrate on Gramian-inspired techniques in Sections 2.3 and 2.4. We show in Example 7 that the deterministic Gramian $\sum_{i=1}^{N} A^{i-1}BB^T\left(A^T\right)^{i-1}$ does not provide a null-controllability metric. We propose a backward stochastic Riccati difference scheme (BSRDS) providing the adequate controllability metric. The link between this BSRDS and null-controllability makes the object of our second main result (Theorem 8). To our best knowledge, these particular schemes are new to the very rich literature on Riccati techniques. Let us just mention that Riccati methods in connection to linear stochastic control problems have been extensively employed in both continuous (cf. [31], etc.) and discrete setting (e.g. [6], [7], etc.). The solvability of the BSRDS and explicit iterative constructions of the solution in particular frameworks make the object of Section 2.5. We study the case of non-random coefficients in Proposition 10. In Proposition 12, we state the solvability of BSRDS with random coefficients in the absence of multiplicative noise. Finally, in Section 2.6, we show that the invariance techniques developed in [2] for Brownian perturbations and adapted to trend-dependent jump-systems in [16] are not suitable in discrete-time. For non-random coefficients, an invariance condition (similar to [16, Criterion 3]) is necessary to achieve null-controllability (cf. Proposition 17). However, it is not sufficient, as shown in Example 19. Concerning the second framework, in absence of multiplicative noise, the continuous-time condition provided in [16, Section 4.2, Criterion 4] is neither necessary (see Example 20) nor sufficient (Example 21).

The aim of Section 3 is to provide a possible application of controllability metrics to biological networks. The mathematical considerations are motivated by the notion of (sub)modularity (see [23, Section 4], [21], etc.) as well as the recent applications to power electronic actuator placement in the preprint [30]. We describe the optimization problems appearing when one works with several (possible) control matrices and wishes to keep controllability features by selecting a minimal dimension of the control space. To end the section, we give a toy model inspired by bacteriophage λ in [17] and analyze different scenarios leading to null-controllability.

Finally, Section 4 gathers the proofs of our mathematical assertions.

2 The Main Concepts and Results

2.1 Controllability and Duality

We begin with recalling the following controllability concepts.

Definition 1 *i. The system (1) is said to be controllable at time $N \geq 1$ if, for every initial data $x \in \mathbb{R}^m$ and every $\mathbb{R}^m-$valued, $\mathcal{F}_N-$measurable square integrable random variable ξ, there exists a predictable control process $u \in \mathcal{P}red$ such that $X_N^{x,u} = \xi$, $\mathbb{P}-a.s.$*

ii. The system (1) is said to be null-controllable at time $N \geq 1$ if the previous property holds true for $\xi = 0$.

iii. The system (1) is said to be approximately controllable at time $N \geq 1$ if, for every initial data $x \in \mathbb{R}^m$ and every $\mathbb{R}^m-$valued, $\mathcal{F}_N-$measurable square integrable random variable ξ and every $\varepsilon > 0$, there exists a predictable control process $u^\varepsilon \in \mathcal{P}red$ such that $\mathbb{E}\left[\left|X_N^{x,u^\varepsilon} - \xi\right|^2\right] \leq \varepsilon$.

iv. The system (1) is said to be approximately null-controllable at time $N \geq 1$ if the previous property holds true for $\xi = 0$.

To the decision process (1), one can associate an adjoint process (or an adjoint couple) as follows. For every $\mathcal{F}_N$-measurable square integrable random variable ξ, we introduce the adjoint couple $\left(Y^{N,\xi}, Z^{N,\xi}\right)$ consisting in an $\mathbb{R}^m$-valued (resp. $\mathbb{R}^{m\times p}$-valued) adapted process by setting

$$(2) \qquad \begin{cases} Y_N^{N,\xi} := \xi, \\ Y_n^{N,\xi} := A_n^T \mathbb{E}\left[Y_{n+1}^{N,\xi}/\mathcal{F}_n\right] + \sum_{i=1}^p C_{i,n}^T Z_n^{N,\xi} \mathbb{E}\left[\langle \Delta M_{n+1}, e_i\rangle \, \Delta M_{n+1}/\mathcal{F}_n\right], \\ \text{where } Y_{n+1}^{N,\xi} = \mathbb{E}\left[Y_{n+1}^{N,\xi}/\mathcal{F}_n\right] + Z_n^{N,\xi}\Delta M_{n+1}, \text{ for all } 0 \leq n \leq N-1. \end{cases}$$

The existence (and uniqueness up to an equivalence) of processes Z satisfying the last property is standard. We refer the interested reader to [3, Corollary 1] or [5, Corollary 3.1.1] and references therein.

The first result of our paper provides the following characterization of controllability.

Theorem 2 *i) The system (1) is approximately null-controllable in time $N > 0$ if and only if every solution $\left(Y_n^{N,\xi}, Z_n^{N,\xi}\right)$ of the scheme (2) satisfying $\mathbb{E}\left[Y_n^{N,\xi}/\mathcal{F}_{n-1}\right] \in \ker\left(B^T\right)$, $\mathbb{P}$-a.s., for all $1 \leq n \leq N$, equally satisfies $Y_0^{N,\xi} = 0$, $\mathbb{P}$-a.s.*

ii) The system (1) is approximately controllable in time $N > 0$ if and only if every solution $\left(Y_n^{N,\xi}, Z_n^{N,\xi}\right)$ of the scheme (2) satisfying $\mathbb{E}\left[Y_n^{N,\xi}/\mathcal{F}_{n-1}\right] \in \ker\left(B^T\right)$, $\mathbb{P}$-a.s., for all $1 \leq n \leq N$, equally satisfies $\mathbb{E}\left[\xi/\mathcal{F}_{n-1}\right] = 0$, $\mathbb{P}$-a.s.

iii) The system (1) is approximately null-controllable in time $N > 0$ if and only if it is (exactly) null-controllable (in time $N > 0$). The necessary and sufficient condition for null-controllability is the existence of some constant $k > 0$ such that

$$(3) \qquad \left|Y_0^{N,\xi}\right|^2 \leq k\mathbb{E}\left[\sum_{n=1}^N \left\langle BB^T\mathbb{E}\left[Y_n^{N,\xi}/\mathcal{F}_{n-1}\right], \mathbb{E}\left[Y_n^{N,\xi}/\mathcal{F}_{n-1}\right]\right\rangle\right],$$

for all $\left(Y_n^{N,\xi}, Z_n^{N,\xi}\right)$ satisfying (2).

The proof is postponed to Section 4. The first two assertions are proven by taking convenient controllability operators and identifying their duals. The third assertion makes use of these duals and of the finite-dimensional setting.

2.2 An Alternative Characterization and an Example

When the linear coefficient A is invertible, we are able to restate the null-controllability criterion given in Theorem 2 iii) by interpreting the adjoint couple as a decision process (where the second component of the couple is an arbitrary predictable control). We also give an example showing that, in the context of discrete processes, null-controllability is, in all generality, strictly weaker that approximate-controllability.

From now on, unless stated otherwise, the matrix $A_n(\omega)$ is assumed to be invertible for $\mathbb{P}-$almost all $\omega \in \Omega$ and all $n \geq 0$. The reader will note that $\left(Y^{N,\xi}, Z^{N,\xi}\right)$ given by (2) can be interpreted in connection to a (forward) decision process by picking $v_{n+1} := Z_n^{N,\xi}$ and setting

$$(4) \qquad \begin{cases} y_0 := Y_0^{N,\xi}, \ y_0^{y_0,v} = y_0, \\ y_{n+1}^{y_0,v} := \left[A_n^T\right]^{-1}\left(y_n^{y_0,v} - \sum_{i=1}^p C_{i,n}^T v_{n+1}\mathbb{E}\left[\langle \Delta M_{n+1}, e_i\rangle \, \Delta M_{n+1}/\mathcal{F}_n\right]\right) + v_{n+1}\Delta M_{n+1}, \end{cases}$$

for all $0 \leq n \leq N-1$.

Remark 3 *1. When A is not invertible, the admissible controls should be such that*

$$y_n^{y_0,v} - \sum_{i=1}^{p} C_{i,n}^T v_{n+1} \mathbb{E}\left[\langle \Delta M_{n+1}, e_i \rangle \, \Delta M_{n+1} / \mathcal{F}_n\right] \in \mathrm{Im}\left(A_n^T\right),$$

where Im *stands for the image of the linear operator. Nevertheless, the connection is still preserved.*

2. The adjoint process is motivated by the duality techniques in the Brownian case (e.g. in [2]). These arguments concern backward stochastic differential equations. The specialization of this concept to discrete-time processes is the notion of backward stochastic difference equation (e.g. [3], [4]). In view of the essential bijection requirement (cf. [3, Theorem 2], [4, Theorem 1.2]), asking for A to be invertible does not appear to be a drawback.

In this framework, the third assertion of Theorem 2 can be interpreted as follows.

Criterion 4 *The system (1) is approximately (and exactly) null-controllable if and only if there exists some $k > 0$ such that for every $y_0 \in \mathbb{R}^m$ and every $\mathbb{F}-$predictable, $\mathbb{R}^{m\times p}$-valued sequence $(v_n)_{1\leq n\leq N}$, one has*

$$|y_0|^2 \leq k\mathbb{E}\left[\sum_{n=0}^{N-1} \left|B^T \mathbb{E}\left[y_{n+1}^{y_0,v}/\mathcal{F}_n\right]\right|^2\right],$$

where $y^{y_0,v}$ is the decision process defined by (4).

In the continuous-time framework, when the controlled linear systems are driven by non-random and homogeneous coefficients (i.e. systems for which A and C are constant matrices independent of n), it has been proven that approximate null-controllability and approximate controllability are equivalent. The reader is referred to [2, Theorem 1.3] (for Brownian setting) and to [15, Theorem 2.2] and [16, Criterion 3] for jump systems. The following example shows that, in the case of discrete-time processes, one can have (exact) null-controllability without having approximate controllability.

Example 5 *To this purpose, let us take $p = 2$ and the transition matrix $Q = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}$. We consider the time horizon $N = 2$, the state space dimension $m = 2$ and the control space dimension $d = 1$. Moreover, we consider*

$$A_n = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \; B = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \; C_{i,n} = \begin{pmatrix} 0 & (-1)^{i+1} \\ 0 & 0 \end{pmatrix}, \text{ for } i \in \{1,2\} \text{ and } n \geq 0.$$

Then, the decision process (1) becomes

$$X_0^{x,u} = x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \; X_1^{x,u} = \begin{pmatrix} (1 + \langle L_1, e_1 - e_2\rangle)\, x_2 \\ x_1 + u_1 \end{pmatrix}, \; X_2^{x,u} = \begin{pmatrix} (x_1 + u_1)\,(1 + \langle L_2, e_1 - e_2\rangle) \\ (1 + \langle L_1, e_1 - e_2\rangle)\, x_2 + u_2 \end{pmatrix}.$$

We consider $u_1 = -x_1$ and $u_2 = -(1 + \langle L_1, e_1 - e_2\rangle)\, x_2$ to conclude that the system is exactly null-controllable in 2 steps. Nevertheless, by considering $\xi := \begin{pmatrix} \langle L_2, e_1 - e_2\rangle \\ 0 \end{pmatrix}$, one has

$$\mathbb{E}\left[\left|X_2^{x,u} - \xi\right|^2\right] \geq \mathbb{E}\left[\left[(x_1 + u_1) + (x_1 + u_1 - 1)\,\langle L_2, e_1 - e_2\rangle\right]^2\right] \geq \frac{1}{2},$$

for all $x \in \mathbb{R}^2$ and all $u_1 \in \mathbb{R}$. The system turns out not to be approximately controllable at time $N = 2$. (In fact, we have proven something stronger: the system is not even exactly terminal controllable; see [25] for a comparison with the Brownian setting).

Remark 6 *To prove null-controllability, one can, alternatively, rely on Criterion 4. For* $y_0 = \begin{pmatrix} y_0^1 \\ y_0^2 \end{pmatrix} \in \mathbb{R}^2$ *and a family of* $\mathbb{F}$*-predictable,* $\mathbb{R}^{2\times 2}$*-valued controls* $v_1 = \begin{pmatrix} v_1^{1,1} & v_1^{1,2} \\ v_1^{2,1} & v_1^{2,2} \end{pmatrix}$, $v_2 = \begin{pmatrix} v_2^{1,1} & v_2^{1,2} \\ v_2^{2,1} & v_2^{2,2} \end{pmatrix}$, *simple (yet fastidious) computations show that*

$$\mathbb{E}\left[\sum_{n=0}^{1}\left|B^T\mathbb{E}\left[y_{n+1}^{y_0,v}/\mathcal{F}_n\right]\right|^2\right] = \left(y_0^1\right)^2 + \mathbb{E}\left[\left(y_0^2 + \left(\langle L_1, e_1 - e_2\rangle - \frac{1}{2}\right)\frac{v_1^{1,1} - v_1^{1,2}}{2}\right)^2\right] \geq \frac{1}{2}\left|y_0\right|^2.$$

One concludes to the exact null-controllability by applying Criterion 4.

2.3 The Deterministic Controllability Metric Is Insufficient

A simple glance at the inequality in Criterion 4 allows one to infer that the right-hand term (i.e. $\mathbb{E}\left[\sum_{n=0}^{N-1}\left|B^T\mathbb{E}\left[y_{n+1}^{y_0,v}/\mathcal{F}_n\right]\right|^2\right]$) should provide a quadratic function of the initial data y_0 which is positive-definite when the initial system is null-controllable. In the deterministic framework ($C = 0$ and non-random, constant A), the controllability criterion is given by the celebrated Kalman condition

$$Rank\left[B\ AB\ A^2B\ ...A^{N-1}B\right] = m.$$

Then, a possible metric would involve the full rank matrix

$$p_0 := \sum_{i=1}^{N} A^{i-1}BB^T\left(A^T\right)^{i-1}\ . \tag{5}$$

In this case, the controllability (pseudo)norm given by $\mathbb{R}^m \ni y_0 \mapsto \left(\langle p_0 y_0, y_0\rangle\right)^{\frac{1}{2}}$ is a norm. So the obvious question one asks oneself is whether the same norm characterizes the stochastic null-controllability. The answer is negative. Unlike the additive case, the presence of multiplicative noise induces a change in the controllability condition. This is not surprising for our reader familiar with the stochastic framework. Indeed, the invariance conditions characterizing approximate null-controllability in [2, Theorem 1.3] or [16, Criterion 3] involve the stochastic component (i.e. C). The following example shows that, in the discrete framework, one may have Kalman's condition and not achieve the null-controllability of the stochastic system.

Example 7 *To this purpose, let us take* $p = 2$ *and the transition matrix* $Q = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}$. *We consider the time horizon* $N = 2$, *the state space dimension* $m = 2$ *and the control space dimension* $d = 1$. *Moreover, we consider*

$$A_n = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},\ C_{i,n} = (-1)^{i+1}A_n,\ B = \begin{pmatrix} 0 \\ 1 \end{pmatrix},\ \textit{for } i \in \{1,2\}\ \textit{and } n \geq 1.$$

We also drop the dependency on n *in* A *and* C. *Then,* $Rank\left[B\ AB\right] = Rank\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = 2$. *However, by taking the initial condition* $x = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, *one gets*

$$X_2^{x,u} = \begin{pmatrix} \left(1 + \langle L_2, e_1 - e_2\rangle\right)\left(1 + \langle L_1, e_1 - e_2\rangle + u_1\right) \\ u_2 \end{pmatrix}.$$

As consequence, for any (predictable) choice of the control u,

$$\mathbb{E}\left[\left|X_2^{x,u}\right|^2\right] = \left(u_1^2 + (2+u_1)^2\right) + \mathbb{E}\left[u_2^2\right] \geq 2.$$

Therefore, independently of the predictable control we use, we are not able to drive the process X from x to 0 even though Kalman's condition is satisfied.

2.4 A Stochastic Controllability Metric

In view of Criterion 4, we associate, to every point y in $\mathbb{R}^m$ the controllability (pseudo)norm

$$\|y\|_{ctrl}^2 := \inf_{(v)_{1\leq n\leq N} \ \mathbb{F}\text{-predictable}} \mathbb{E}\left[\sum_{n=0}^{N-1} \left|B^T \mathbb{E}\left[y_{n+1}^{y_0,v}/\mathcal{F}_n\right]\right|^2\right]. \tag{6}$$

The previous example shows that the associated metric is a stochastic one and, in general, it does not reduce to the deterministic Gramian. Nevertheless, one would very much like to have something which is close to the p_0 matrix in (5). In this section, we thrive to provide a (pseudo)metric which is more explicit than (6).

To this purpose, let us analyze the following matrix scheme. We set, for $\varepsilon > 0$, $P_N^\varepsilon = 0 \in \mathbb{R}^{m\times m}$ and proceed by writing, for all $0 \leq n \leq N-1$,

$$P_{n+1}^\varepsilon = \mathbb{E}\left[P_{n+1}^\varepsilon/\mathcal{F}_n\right] + Q_n^\varepsilon \, diag\left(\Delta M_{n+1}\right),$$

where, by convention,

$$diag\left(\Delta M_{n+1}\right) = \begin{pmatrix} \Delta M_{n+1} & 0 & ... & 0 \\ 0 & \Delta M_{n+1} & ... & 0 \\ ... & ... & ... & \\ 0 & 0 & ... & \Delta M_{n+1} \end{pmatrix} \in \mathbb{R}^{mp\times m}. \tag{7}$$

The existence and uniqueness of such $Q_n^\varepsilon \in \mathbb{R}^{m\times mp}$ is obtained by applying the martingale representation theorem (see, for example [3, Corollary 1] or [5, Corollary 3.1.1] and references therein) to the columns of P_{n+1}^ε. We proceed by setting

$$P_n^\varepsilon = A_n^{-1}\left(\mathbb{E}\left[P_{n+1}^\varepsilon/\mathcal{F}_n\right] + BB^T\right)\left[A_n^T\right]^{-1} - \alpha_{n,\varepsilon}^T \eta_{n,\varepsilon}^{-1} \alpha_{n,\varepsilon}, \tag{8}$$

where $\alpha_{n,\varepsilon} = \begin{pmatrix} \alpha_{n,\varepsilon}^1 \\ ... \\ \alpha_{n,\varepsilon}^p \end{pmatrix} \in \mathbb{R}^{mp\times m}$, $\eta_{n,\varepsilon} = \begin{pmatrix} \eta_{n,\varepsilon}^{1,1} & \eta_{n,\varepsilon}^{1,2} & ... & \eta_{n,\varepsilon}^{1,p} \\ \eta_{n,\varepsilon}^{2,1} & \eta_{n,\varepsilon}^{2,2} & ... & \eta_{n,\varepsilon}^{2,p} \\ ... & ... & ... & ... \\ \eta_{n,\varepsilon}^{p,1} & \eta_{n,\varepsilon}^{p,2} & ... & \eta_{n,\varepsilon}^{p,p} \end{pmatrix} \in \mathbb{R}^{mp\times mp}$ are given by

$$\begin{aligned} \alpha_{n,\varepsilon}^j := & -Q_n^\varepsilon \mathbb{E}\left[\left\langle \Delta M_{n+1}, e_j\right\rangle diag\left(\Delta M_{n+1}\right)/\mathcal{F}_n\right]\left[A_n^T\right]^{-1} \\ & + \sum_{1\leq i\leq p} \mathbb{E}\left[\left\langle \Delta M_{n+1}, e_j\right\rangle \left\langle \Delta M_{n+1}, e_i\right\rangle/\mathcal{F}_n\right] C_{i,n} A_n^{-1}\left(\mathbb{E}\left[P_{n+1}^\varepsilon/\mathcal{F}_n\right] + BB^T\right)\left[A_n^T\right]^{-1} \text{ and} \end{aligned}$$

$$
\begin{aligned}
&\eta^{j,k}_{n,\varepsilon} := \\
&\varepsilon\delta_{j,k} I_{m\times m} + \frac{1}{2} Q^{\varepsilon}_n \mathbb{E}\left[\langle \Delta M_{n+1}, e_k\rangle \langle \Delta M_{n+1}, e_j\rangle \, diag\left(\Delta M_{n+1}\right)/\mathcal{F}_n\right] \\
&+ \frac{1}{2}\mathbb{E}\left[\langle \Delta M_{n+1}, e_k\rangle \langle \Delta M_{n+1}, e_j\rangle \left(diag\left(\Delta M_{n+1}\right)\right)^T/\mathcal{F}_n\right] \left(Q^{\varepsilon}_n\right)^T \\
&- \frac{1}{2}\sum_{1\le i\le p} Q^{\varepsilon}_n \mathbb{E}\left[\langle \Delta M_{n+1}, e_i\rangle \langle \Delta M_{n+1}, e_j\rangle/\mathcal{F}_n\right]\mathbb{E}\left[\langle \Delta M_{n+1}, e_k\rangle \, diag\left(\Delta M_{n+1}\right)/\mathcal{F}_n\right]\left[A^T_n\right]^{-1} C^T_{i,n} \\
&- \frac{1}{2}\sum_{1\le i\le p} C_{i,n} A^{-1}_n \mathbb{E}\left[\langle \Delta M_{n+1}, e_i\rangle \langle \Delta M_{n+1}, e_k\rangle/\mathcal{F}_n\right]\mathbb{E}\left[\langle \Delta M_{n+1}, e_j\rangle \left(diag\left(\Delta M_{n+1}\right)\right)^T/\mathcal{F}_n\right]\left(Q^{\varepsilon}_n\right)^T \\
&+ \mathbb{E}\left[\langle \Delta M_{n+1}, e_k\rangle \langle \Delta M_{n+1}, e_j\rangle/\mathcal{F}_n\right]\mathbb{E}\left[P^{\varepsilon}_{n+1}/\mathcal{F}_n\right] \\
&+ \sum_{1\le i,i'\le p}\left(\begin{array}{c}\mathbb{E}\left[\langle \Delta M_{n+1}, e_j\rangle \langle \Delta M_{n+1}, e_i\rangle/\mathcal{F}_n\right]\times\mathbb{E}\left[\langle \Delta M_{n+1}, e_k\rangle \langle \Delta M_{n+1}, e_{i'}\rangle/\mathcal{F}_n\right]\times \\ \times C_{i',n} A^{-1}_n\left(\mathbb{E}\left[P^{\varepsilon}_{n+1}/\mathcal{F}_n\right] + BB^T\right)\left[A^T_n\right]^{-1} C^T_{i,n}\end{array}\right),
\end{aligned}
$$

for all $1 \le j,k \le p$. Here, $\delta_{j,k}$ stands for the classical Kronecker delta (1, if $j=k$ and 0, otherwise). While it is clear that $\eta_{n,\varepsilon}$ is symmetric , it is (a lot) less obvious to ask for $\eta_{n,\varepsilon}$ to be positive. We will show in some particular cases that this stochastic Riccati-type difference equation is solvable and provide explicit construction for P and Q. For the time being, let us assume that, for every $\varepsilon > 0$, such a solution exists. The second main result of our paper is the following characterization of the null-controllability.

Theorem 8 *i. The system (1) is (approximately) null-controllable if and only if*

$$\liminf_{\varepsilon\to 0+} P^{\varepsilon}_0 \text{ is a positive-definite, symmetric matrix.}$$

ii. The controllability (pseudo)norm given by (6) satisfies

$$\|y_0\|^2_{ctrl} = \liminf_{\varepsilon\to 0}\langle P^{\varepsilon}_0 y_0, y_0\rangle.$$

The proof is postponed to Section 4. The construction of P^{ε} is tailor-made to infer a recurrence on the terms $\langle P^{\varepsilon}_n y^{y_0,v}_n, y^{y_0,v}_n\rangle$. To conclude, one applies Criterion 4.

Remark 9 *i. This result is the discrete-time version of [29, Theorem 3.4].*

ii. A simple look at the proof (see (24)) shows that

$$\langle P^{\varepsilon}_0 y_0, y_0\rangle = \inf_{(v_n)_{1\le n\le N}\ \mathbb{F}\text{-predictable}}\left(\varepsilon\sum_{n=0}^{N-1}\mathbb{E}\left[|v_{n+1}|^2\right] + \mathbb{E}\left[\sum_{n=0}^{N-1}\left|B^T\mathbb{E}\left[y^{y_0,v}_{n+1}/\mathcal{F}_n\right]\right|^2\right]\right)$$

and the optimal control realizing this minimum is given in feedback form by

$$v^{opt}_{n+1} = \eta^{-1}_{n,\varepsilon}\delta_{n,\varepsilon} y^{y_0,v^{opt}}_n.$$

Nevertheless, δ might not be a Markov process and, hence, a fortiori, neither would v^{opt}.

2.5 Solvability of the Backward Stochastic Riccati Difference Scheme (BSRDS)

The aim of this subsection is to provide two simple cases in which the backward stochastic Riccati scheme admits explicit solutions. One of the simplest frameworks for our trend component is the one in which the martingale is generated by independent and identically distributed variables. In other words, we assume L_{n+1} to be independent of $\mathcal{F}_n$ for all $n \ge 0$ and L_n has the same law as L_0. Then $(\langle L_n, e_i\rangle)_{n\ge 1}$ are independent Bernoulli distributed with some parameter $q_i > 0$ (independent of n) and such that $\sum_{i=1}^p q_i = 1$.

The first setting is when A and C consist in sequences of non-random matrices. In other words, the randomness may only come from the martingale induced by the trend component L. In this case, we get the following result of solvability of the BSRDS.

Proposition 10 (non-random coefficients case) *We assume that L_n are independent, identically distributed random variables on $\{e_1, e_2, ..., e_p\}$ and denote by*

$$q_i = \mathbb{P}(L_1 = e_i) > 0, \text{ for every } 1 \leq i \leq p.$$

Moreover, we assume that

$$A_n,\, C_n \in \mathbb{R}^{m\times m}, \text{ for all } n \geq 0$$

are sequences of (non-random) matrices. Then, for every $\varepsilon > 0$, the BSRDS (8) admits a (unique) solution given by a positive-semidefinite sequence $(P_n^\varepsilon)_{0\leq n\leq N}$ and $Q^\varepsilon = 0$. This solution is given by

$$(9)\qquad \begin{cases} P_N^\varepsilon = 0, \\ \alpha_{n,\varepsilon} = \mathcal{C}_n A_n^{-1}\left(P_{n+1}^\varepsilon + BB^T\right)\left[A_n^T\right]^{-1}, \\ \eta_{n,\varepsilon} = \varepsilon I_{mp\times mp} + \left(q_j\left(\delta_{j,k} - q_k\right)P_{n+1}^\varepsilon\right)_{1\leq j,k\leq p} + \mathcal{C}_n A_n^{-1}\left(P_{n+1}^\varepsilon + BB^T\right)\left[\mathcal{C}_n A_n^{-1}\right]^T, \\ P_n^\varepsilon = A_n^{-1}\left(P_{n+1}^\varepsilon + BB^T\right)\left[A_n^T\right]^{-1} - \alpha_{n,\varepsilon}^T \eta_{n,\varepsilon}^{-1}\alpha_{n,\varepsilon}, \end{cases}$$

where $\mathcal{C}_n := \begin{bmatrix} \sum_{l=1}^p \left(\delta_{1,l} - q_1\right) q_l C_{l,n} \\ \sum_{l=1}^p \left(\delta_{2,l} - q_2\right) q_l C_{l,n} \\ \cdots \\ \sum_{l=1}^p \left(\delta_{p,l} - q_p\right) q_l C_{l,n} \end{bmatrix}.$

The proof follows by (descending) induction and is postponed to Section 4.

Remark 11 *i. We emphasize that we deal here with a difference equation and not with an algebraic Riccati equation and this is why one does not need further conditions on the Popov matrix.*

ii. The Riccati difference equation given by (9) is obviously deterministic. Then, a glance at the optimal control in Remark 9 shows that $v_{n+1}^{opt} = \eta_{n,\varepsilon}^{-1}\delta_{n,\varepsilon} y_n^{y_0, v^{opt}}$ is not only predictable but a deterministic function of time n and the state of the process y_n. Therefore, the infimum in $\langle P_0^\varepsilon y_0, y_0\rangle$ can be taken over open-loop control strategies. Of course, similar assertions hold true for the controllability (pseudo)norm. As a by-product the process y constructed with open-loop controls is Markovian.

The second case in which solving the backward stochastic Riccati difference equation is reduced to deterministic arguments is when $C = 0$. Under this assumption, the BSRDS becomes

$$(10)\qquad \begin{cases} P_n^\varepsilon := A_n^{-1}\left(\mathbb{E}\left[P_{n+1}^\varepsilon/\mathcal{F}_n\right] + BB^T\right)\left[A_n^T\right]^{-1} - \alpha_{n,\varepsilon}^T \eta_{n,\varepsilon}^{-1}\alpha_{n,\varepsilon}, \\ \alpha_{n,\varepsilon}^j := -q_j Q_n^\varepsilon diag\left(\sum_{k=1}^p \left(\delta_{j,k} - q_k\right) e_k\right)\left[A_n^T\right]^{-1}, \\ \eta_{n,\varepsilon}^{j,k} := \varepsilon\delta_{j,k} I_{m\times m} + \left(\delta_{j,k} - q_k\right) q_j \mathbb{E}\left[P_{n+1}^\varepsilon/\mathcal{F}_n\right] + \frac{1}{2} Q_n^\varepsilon diag\left(\sum_{l=1}^p m_{j,k,l} e_l\right) \\ + \frac{1}{2}\left(diag\left(\sum_{l=1}^n m_{j,k,l} e_l\right)\right)^T \left(Q_n^\varepsilon\right)^T. \end{cases}$$

for all $1 \leq j, k \leq p$. Here,

$$m_{j,k,l} = q_l\left(q_j - \delta_{j,l}\right)\left(q_k - \delta_{k,l}\right) - q_l\left(\delta_{j,k} - q_j\right) q_k.$$

Let us recall that the *diag* matrices are of type $\mathbb{R}^{mp\times m}$ and are constructed as in (7).

The main result in this framework gives the solvability of the BSRDS (10).

Proposition 12 (the case without multiplicative noise) *We assume that L_n are independent, identically distributed random variables on $\{e_1, e_2, ..., e_p\}$ and denote by*

$$q_i = \mathbb{P}(L_1 = e_i) > 0, \text{ for every } 1 \leq i \leq p.$$

Moreover, we assume that $A_n = A(n, L_n)$ where A is some $\mathbb{R}^{m\times m}$-valued deterministic function of time and trend.

Then, for every $\varepsilon > 0$ there exist two sequences of positive-semidefinite matrices $(p_n^\varepsilon)_{0\leq n\leq N}$ and $(q_n^\varepsilon)_{0\leq n\leq N}$ such that

$$p_n^\varepsilon \geq q_n^\varepsilon, \text{ for all } 0 \leq n \leq N \tag{11}$$

and the couple $(P^\varepsilon, Q) \in \mathbb{R}^{m\times m} \times \mathbb{R}^{m\times mp}$ defined by $P_N^\varepsilon = 0$ and $Q_{N-1} = 0$

$$P_n^\varepsilon = A^{-1}(n, L_n)\left(p_{n+1}^\varepsilon + BB^T - q_{n+1}^\varepsilon\right)\left(A^{-1}(n, L_n)\right)^T, \text{ for all } 0 \leq n \leq N-1 \tag{12}$$

and

$$\begin{aligned} Q_{n-1} &= \left[Q_{n-1,1}^{\cdot,\cdot}, Q_{n-1,2}^{\cdot,\cdot}, ..., Q_{n-1,m}^{\cdot,\cdot}\right] \in \mathbb{R}^{m\times p} \times \mathbb{R}^{m\times p} \times ... \times \mathbb{R}^{m\times p}, \text{ where} \\ &\left[Q_{n-1,1}^{\cdot,l}, Q_{n-1,2}^{\cdot,l}, ..., Q_{n-1,m}^{\cdot,l}\right] = A^{-1}(n, e_l)\left(p_{n+1}^\varepsilon + BB^T - q_{n+1}^\varepsilon\right)\left(A^{-1}(n, e_l)\right)^T, \end{aligned} \tag{13}$$

for all $1 \leq l \leq p$, $1 \leq n \leq N-1$ is the solution of (10).

Remark 13 *i. As we will see in the proof, p and q are explicitly given by setting*

$$\begin{cases} p_{n+1}^\varepsilon := \sum_{l=1}^{p} q_l A^{-1}(n+1, e_l)\left(p_{n+2}^\varepsilon + BB^T - q_{n+2}^\varepsilon\right)\left(A^{-1}(n+1, e_l)\right)^T, \\ q_{n+1}^\varepsilon = \overline{\alpha}_{n,\varepsilon}^T \eta_{n,\varepsilon}^{-1} \overline{\alpha}_{n,\varepsilon}, \end{cases} \tag{14}$$

where

$$\begin{cases} \overline{\alpha}_{n,\varepsilon}^j = \left[\sum_{l=1}^{p} q_l(\delta_{j,l} - q_j) A^{-1}(n+1, e_l)\left(p_{n+2}^\varepsilon + BB^T - q_{n+2}^\varepsilon\right)\left(A^{-1}(n+1, e_l)\right)^T\right], \\ \alpha_{n,\varepsilon}^j = -\overline{\alpha}_{n,\varepsilon}^j \left[A_n^T\right]^{-1}. \\ \eta_{n,\varepsilon}^{j,k} = \varepsilon\delta_{j,k} I_{m\times m} \\ \qquad + \sum_{l=1}^{p} q_l(q_j - \delta_{j,l})(q_k - \delta_{k,l}) A^{-1}(n+1, e_l)\left(p_{n+2}^\varepsilon + BB^T - q_{n+2}^\varepsilon\right)\left(A^{-1}(n+1, e_l)\right)^T, \end{cases} \tag{15}$$

for all $1 \leq j, k \leq p$.

ii. When one further assumes that A is non-random, the elements $Q_{n-1,k}^{\cdot,l}$ are independent of l. Hence,

$$Q_{n-1,k}^{\cdot,\cdot}\Delta M_{n+1} = Q_{n-1,k}^{\cdot,1}\sum_{l=1}^{p}(\langle L_{n+1}, e_l\rangle - q_l) = 0_{m\times 1} = 0_{m\times p}\Delta M_{n+1},$$

i.e. Q_{n-1} is equivalent (see, for example [3, Definition 2]) to $0_{m\times mp}$. This is consistent with the results in our non-random framework.

To end this subsection, let us take a look at the case when $C(\cdot) = 0$ and A_n is a non-random matrix. Using the second point of the previous remark, one gets $\alpha = 0_{m\times mp}$ and $\eta_{n,\varepsilon} = \varepsilon I_{mp\times mp}$. Hence, one only has to solve the following (deterministic, ε-independent scheme) :

$$p_n = A_n^{-1}\left(p_{n+1} + BB^T\right)\left(A_n^T\right)^{-1}, \; p_N = 0_{m\times m}.$$

In this framework, we get the following criterion.

Criterion 14 *Let us assume that $C(\cdot)=0$ and A_n is a non-random matrix, for all $n\geq 1$. Then, the system (1) is null-controllable in time $N>0$ if and only if the matrix*

$$p_0^N=\sum_{n=0}^{N-1}\left[\left(\prod_{k=0}^{n}A_k^{-1}\right)BB^T\left(\prod_{k=0}^{n}A_k^{-1}\right)^T\right]$$

has full rank.

Remark 15 *If A does not depend on n, p_0 is of full rank if and only if $A^N p_0\left(A^T\right)^N$ is of full rank and one gets the classical condition $p_0=\sum_{n=0}^{N-1}A^nBB^T\left(A^T\right)^n$ is of full rank.*

2.6 When Continuous-time Intuition Fails to Work

As we have seen in Example 5, the null-controllability metric is given by a strictly weaker condition than that of controllability. The reader acquainted with the continuous-time characterizations of approximate controllability ([2] or [14] for Brownian perturbations, [15] or [16] for continuous-time jump processes) may wonder whether alternative approaches based on invariance concepts can be adapted to this discrete framework. The aim of this section is to compare our approach with the algebraic conditions given in [16] for continuous-time processes presenting a similar construction. We will consider both the non-random coefficients setting and the behavior of the system lacking multiplicative noise.

In the case of non-random coefficients, the method developed in [2] for Brownian perturbations and adapted to trend-dependent jump-systems in [16] consists in obtaining invariance criteria starting from (4). We will prove that the analogous condition is still necessary in order to have null-controllability. Nevertheless, this condition is strictly weaker than the one exhibited in Theorem 8 (see Example 19). Concerning the second framework, in absence of multiplicative noise, the continuous-time condition provided in [16, Section 4.2, Criterion 4] is neither necessary (see Example 20) nor sufficient (Example 21).

Throughout the subsection, we assume that L_n are independent, identically distributed random variables on $\{e_1,e_2,...,e_p\}$ and denote by $q_i=\mathbb{P}(L_1=e_i)>0$, for every $1\leq i\leq p$.

2.6.1 The Non-Random Coefficients Case

To simplify the arguments, we concentrate on the time-homogeneous framework (i.e. $A_n=A\in\mathbb{R}^{m\times m}$, $C_{i,n}=C_i\in\mathbb{R}^{m\times m}$, for all $n\geq 0$). In this setting, the dual decision process (4) becomes

$$y_{n+1}^{y_0,v}:=\left[A^T\right]^{-1}\left(y_n^{y_0,v}-\sum_{l=1}^{p}\mathcal{C}^T(l)\,v_{n+1}e_l\right)+\sum_{l=1}^{p}\langle\Delta M_{n+1},e_l\rangle\,v_{n+1}e_l,\ y_0^{y_0,v}=y_0,$$

where

$$\mathcal{C}(j):=\sum_{k=1}^{p}(\delta_{j,k}-q_j)\,q_kC_k, \tag{16}$$

for every $1\leq j\leq p$.

In [16], the study of controllability properties is conducted using some invariance properties with respect to the dual decision process. We recall the following invariance notions.

Definition 16 *We consider a linear operator $\mathcal{A}\in\mathbb{R}^{m\times m}$ and a family $(\mathcal{C}_i)_{1\leq i\leq t}\subset\mathbb{R}^{m\times m}$.*

(a) A set $V\subset\mathbb{R}^m$ is said to be $\mathcal{A}$- invariant if $\mathcal{A}V\subset V$.

(b) A set $V\subset\mathbb{R}^m$ is said to be $(\mathcal{A};\mathcal{C})$- invariant if $\mathcal{A}V\subset V+\mathrm{Im}\,\mathcal{C}_1+\mathrm{Im}\,\mathcal{C}_2+...+\mathrm{Im}\,\mathcal{C}_t$, where Im *stands for the image of the linear operators.*

(c) A set $V \subset \mathbb{R}^m$ is said to be $(\mathcal{A};\mathcal{C})$- strictly invariant if

$$\mathcal{A}V \subset V + \mathrm{Im}\left(\mathcal{C}_1\Pi_V\right) + \mathrm{Im}\left(\mathcal{C}_2\Pi_V\right) + ... + \mathrm{Im}\left(\mathcal{C}_p\Pi_V\right),$$

where Π_V denotes the orthogonal projection onto V.

(d) A set $V \subset \mathbb{R}^n$ is said to be feedback $(\mathcal{A};\mathcal{C})$- invariant if there exists a family of linear operators $(\mathcal{F}_i)_{1\leq i\leq t} \subset \mathbb{R}^{m\times m}$ such that $\left(\mathcal{A} + \sum_{i=1}^t \mathcal{C}_i\mathcal{F}_i\right) V \subset V$ (i.e. V is $\mathcal{A}+\sum_{i=1}^t \mathcal{C}_i\mathcal{F}_i$- invariant).

The following condition is necessary in order to have null-controllability.

Proposition 17 *If the system (1) is null controllable, then*

(N1) *the largest subspace of the kernel $\ker\left(B^T\right)$ which is $\left(\left[A^T\right]^{-1}; \left(\mathcal{C}\left(1\right)A^{-1}\right)^T, ..., \left(\mathcal{C}\left(p\right)A^{-1}\right)^T\right)$-strictly invariant is reduced to $\{0\}$.*

The reader should compare this with [16, Section 4.1, Criterion 3]. The proof of this result is very similar to that of [16, Section 3.1.2, Proposition 2] and is postponed to Section 4.

Remark 18 *i. Both the assertion and the proof can be extended to non-homogeneous systems providing the complete analogue of [16, Section 3.1.2, Proposition 2].*

ii. Much as in the continuous-time framework (see [16, Section 4.1, Criterion 3]), one can prove the equivalence between the following

(a) condition (N1) holds true;

(b) every solution of (4) satisfying $B^T y_n^{y_0,v} = 0$, $\mathbb{P}-a.s.$ for all $n \geq 0$ is such that $y_0 = 0$.

Nevertheless, unlike the continuous-time framework, the condition (N1) is weaker than the null-controllability property. Let us, once again, look at the following example.

Example 19 *We consider $p = 2$ and the transition matrix $Q = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}$, the horizon $N = 2$, the state space dimension $m = 2$ and the control space dimension $d = 1$. Moreover, we consider*

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},\ C_i = (-1)^{i+1} A,\ B = \begin{pmatrix} 0 \\ 1 \end{pmatrix},\ \text{for } i \in \{1,2\}.$$

Then, according to (16), $\mathcal{C}\left(i\right) = \frac{(-1)^{i+1}}{2}A$, for $i \in \{1,2\}$. Moreover, $\ker\left(B^T\right) = \left\{ \begin{pmatrix} x \\ 0 \end{pmatrix},\ x \in \mathbb{R} \right\}$. If $x \in \mathbb{R}$ is such that, for some $x', x'' \in \mathbb{R}$,

$$\left[A^T\right]^{-1} \begin{pmatrix} x \\ 0 \end{pmatrix} + \left(\mathcal{C}\left(1\right)A^{-1}\right)^T \begin{pmatrix} x' \\ 0 \end{pmatrix} + \left(\mathcal{C}\left(2\right)A^{-1}\right)^T \begin{pmatrix} x'' \\ 0 \end{pmatrix} = \begin{pmatrix} \frac{x'-x''}{2} \\ x \end{pmatrix} \in \ker\left(B^T\right),$$

then it follows that $x = 0$. This means that condition (N1) is satisfied. However, as shown in Example 7, the decision system driven by A, B and C is not null-controllable. Thus, in all generality, for discrete-time processes, the condition (N1) does not guarantee null-controllability.

2.6.2 The Case Without Multiplicative Noise (C=0)

In the continuous-time framework, the necessary and sufficient condition for approximate null-controllability of continuous switch systems (equally when $C = 0$, see [16, Section 4.2, Criterion 4]) reads

(17) $$(A_n, B) \text{ is controllable for all } n.$$

Unlike the continuous-time setting, we will see that this condition is neither necessary nor sufficient.

We begin with an example showing that (17) may hold without implying the null-controllability of the discrete system.

Example 20 *We consider the state space dimension $m = 3$ and the control space dimension $d = 1$. Moreover, we consider*

$$A_{2n+1} = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \; A_{2n} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}, \; B = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}.$$

It is obvious that Kalman's condition is satisfied for both A_{2n} and A_{2n+1}. However, by computing p_0^N (see Criterion 14), one gets

$$p_0^N = \sum_{n=0}^{N-1} \left[\left(\prod_{k=0}^{n} A_k^{-1} \right) BB^T \left(\prod_{k=0}^{n} A_k^{-1} \right)^T \right] = \begin{pmatrix} \left\lfloor \frac{N}{2} \right\rfloor & 0 & 0 \\ 0 & \left\lfloor \frac{N+1}{2} \right\rfloor & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

which is obviously not invertible for any $N \geq 1$. Here, $\lfloor \cdot \rfloor$ denotes the largest integer that does not exceed the argument (floor function).

But null-controllability may hold without having (17) for any A_n.

Example 21 *We consider the state space dimension $m = 3$ and the control space dimension $d = 1$. Moreover, we consider*

$$A_{2n+1} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \; A_{2n} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \; \text{for } n \geq 0, \; B = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}.$$

Then Rank $\left[B \; A_{2n+1}B \; A_{2n+1}^2 B\right] = 2$ and Rank $\left[B \; A_{2n}B \; A_{2n}^2 B\right] = 1$. Nevertheless, for $N = 4$,

$$p_0^4 = \sum_{n=0}^{3} \left[\left(\prod_{k=0}^{n} A_k^{-1} \right) BB^T \left(\prod_{k=0}^{n} A_k^{-1} \right)^T \right] = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

is of full rank such that, using Criterion 14, the system is null-controllable at time $N = 4$. The reader may equally want to note that the controllability condition does not hold true at $N' = 3 = m$, the dimension of our state space.

3 A Minimal Intervention-Targeted Application in Biological Networks

The aim of this section is to provide a possible application of the previous mathematical tools to biological networks. The mathematical considerations are motivated by the notion of (sub)modularity (see [23, Section 4], [21], etc.) as well as the recent applications to power electronic actuator placement in the preprint [30]. We describe the optimization problems appearing when one works with several (possible) control matrices. To end the section, we give a toy model inspired by bacteriophage λ in [17].

3.1 Intervention Scenarios

Up until now, the control matrix B has been fixed. We are going to envisage some scenarios translated by several control matrices. Each fundamental matrix allows one-dimensional controls. By putting together some of these matrices (say d), we get an $m \times d$ control matrix taking into account d-dimensional controls. Of course, in the case in which several configurations allow null

controllability, it would be interesting if we were able to find a minimal d (lowest dimension for control processes) giving the null controllability.

We begin with noting that the (pseudo)norm (6) will explicitly depend on the control matrix B and will be denoted by $\|\cdot\|_{ctrl,B}$. Similar, whenever P_0^{ε} given by (8) exists, it is written as $P_0^{\varepsilon}(B)$. We define

$$\|B\|_{ctrl}^{spec} := \inf_{y\neq 0} \frac{\|y\|_{ctrl,B}}{\|y\|} \text{ and } \|B\|_{ctrl}^{rank} := Rank\left(\liminf_{\varepsilon\to 0} P_0^{\varepsilon}(B)\right)$$

It is obvious that the system (1) is controllable using B iff $\|B\|_{ctrl}^{spec} > 0$ or, equivalently, iff $\|B\|_{ctrl}^{rank} = m$.

A basic intervention scenario is a column vector $b \subset \mathbb{R}^m$ allowing one-dimensional controls and specifying the weight of this control in the state component. In other words, we consider the system controlled by $B = b$ and with $d = 1$ in (1). Given a family of $r \in \mathbb{N}^*$ intervention scenarios $\{b_1, b_2, ..., b_r\} \subset \mathbb{R}^m$, for every $\mathcal{I} = \{i_1, i_2, ...i_{|\mathcal{I}|}\} \subset \{1, ...r\}$ one constructs $B(\mathcal{I}) = \left[b_{i_1}, ..., b_{i_{|\mathcal{I}|}}\right]$. We introduce the following definition.

Definition 22 *1) A set $\mathcal{I}$ is called minimal spectral-efficient intervention if the following assertions hold simultaneously:*

(i) one has $\|B(\mathcal{I})\|_{ctrl}^{spec} > 0$;

(ii) for every $\mathcal{J} \subset \{1, ...r\}$ such that $|\mathcal{J}| < |\mathcal{I}|$, one has $\|B(\mathcal{J})\|_{ctrl}^{spec} = 0$;

(iii) for every $\mathcal{J} \subset \{1, ...r\}$ such that $|\mathcal{J}| = |\mathcal{I}|$, one has $\|B(\mathcal{J})\|_{ctrl}^{spec} \leq \|B(\mathcal{I})\|_{ctrl}^{spec}$.

2) A set $\mathcal{I}$ is called minimal rank-efficient intervention if the following assertions hold simultaneously:

(i) one has $\|B(\mathcal{I})\|_{ctrl}^{rank} = m$;

(ii) for every $\mathcal{J} \subset \{1, ...r\}$ such that $|\mathcal{J}| < |\mathcal{I}|$, one has $\|B(\mathcal{J})\|_{ctrl}^{rank} < m$.

The reader is invited to note that any minimal spectral-efficient intervention is also minimal rank-efficient. A condition of type (iii) has no meaning for the rank, being trivially satisfied as soon as $\mathcal{I}$ is minimal rank-efficient.

To find minimal efficient intervention, one has to solve at most r set-function optimization problems of type

$$\max_{\substack{\mathcal{I}\subset\{1,...r\}\\ |\mathcal{I}|=k}} \|B(\mathcal{I})\|_{ctrl}, \ 1 \leq k \leq r,$$

where $\|\cdot\|_{ctrl}$ denotes either $\|\cdot\|_{ctrl}^{spec}$ or $\|\cdot\|_{ctrl}^{rank}$. It is obvious that

$$\min\left\{k : 1 \leq k \leq r, \max_{\substack{\mathcal{I}\subset\{1,...r\}\\ |\mathcal{I}|=k}} \|B(\mathcal{I})\|_{ctrl}^{rank} = m\right\} = \min\left\{k : 1 \leq k \leq r, \max_{\substack{\mathcal{I}\subset\{1,...r\}\\ |\mathcal{I}|=k}} \|B(\mathcal{I})\|_{ctrl}^{spec} > 0\right\}.$$

At this point, one may note that working with minimal spectral-efficient interventions gives more information and may wonder why we have introduced the two concept. It turns out that, although both set functions are non-decreasing, rank-based functions have another useful (submodularity) property (cf. [23], [21]; see also [30]). Let us recall the definition of this concept.

Definition 23 *Given a finite set S, a real-valued function $f : 2^S \longrightarrow \mathbb{R}$ is said to be submodular if*

$$f(S_1 \cap S_2) + f(S_1 \cup S_2) \leq f(S_1) + f(S_2),$$

for all subsets $S_1, S_2 \subset S$.

According to [23], submodularity is "a combinatorial analogue of concavity" in the sense that if the cost functional is submodular, although the problem is NP-hard, a greedy approach provides good results. The paper [23, Section 4] equally provides greedy heuristics as well as relative error bounds concerning the greedy solution.

A glance at [21, Example 1.2] shows that rank-based set functions are submodular. It turns out that, although it provides more information, $\|\cdot\|_{ctrl}^{spec}$ does not, in all generality, provide a submodular application. For an example in this direction, the reader is referred to [30, III.F].

To sum up the considerations made so far, one should begin with solving the optimization problem using $\|\cdot\|_{ctrl}^{rank}$ which is faster (using greedy heuristic as in [23, Section 4]). This will provide a minimal k for which efficient interventions exist. Then, for this particular k, one may work with $\|\cdot\|_{ctrl}^{spec}$.

3.2 Hasty et al.-Inspired Toy Model

Description We start with a toy example concerning the temperate λ virus. The authors of [17] propose a genetic applet consisting in a mutant system in which two operator sites (OR2 and OR3) are present. The gene cI expresses repressor (CI), which dimerizes and binds to the DNA as a transcription factor in one of the two available sites. The site OR2 leads to enhanced transcription, while OR3 represses transcription. We let R_1 stand for the repressor, R_2 for the dimer, D for the DNA promoter site, DR_2 for the binding to the OR2 site, DR_2^* for the binding to the OR3 site and DR_2R_2 for the binding to both sites. We also denote by P the RNA polymerase concentration and by r the number of repressors per mRNA transcript. The capital letters K_i, $1 \leq i \leq 4$ for the reversible reactions correspond to couples of direct/reverse speed functions k_i, k_{-i}, while K_t and K_d only to direct speed functions k_t and k_d.The actual system of biochemical reactions that govern the genetic applet is given by

$$\begin{cases} 2R_1 \overset{K_1}{\rightleftarrows} R_2,\ D\,(+R_2) \overset{K_2}{\rightleftarrows} DR_2,\ D\,(+R_2) \overset{K_3}{\rightleftarrows} DR_2^*, \\ DR_2\,(+R_2) \overset{K_4}{\rightleftarrows} DR_2R_2,\ DR_2 + P \overset{K_t}{\to} DR_2 + P + rR_1,\ R_1 \overset{K_d}{\to}. \end{cases}$$

The Trend Component L We consider the trend component given by the DNA mechanism of the host E-Coli

$$(D, DR_2, DR_2^*, DR_2R_2)^T \text{ which belongs to the basis } \mathcal{B} \subset \mathbb{R}^4.$$

All the reactions concerning at least one of these components is considered to belong to the trend mechanism. The remaining reactions

$$2R_1 \overset{K_1}{\rightleftarrows} R_2, R_1 \overset{K_d}{\to}$$

will be employed to describe the repressor's updating. To simplify the arguments (recall that this is a toy model), we consider that all the speeds in the trend mechanism are unitary ($k_{\pm 2} = k_{\pm 3} = k_{\pm 4} = k_t = 1$). Whenever the system is at position e_1 (unoccupied host DNA), two reactions are possible $D \overset{k_2}{\to} DR_2$, respectively $D \overset{k_3}{\to} DR_2^*$. We consider that transition probability is proportional to the speed of reaction (similar to [10]) to get

$$\mathbb{P}\,(L_{n+1} = e_2 / L_n = e_1) = \frac{k_2}{k_2 + k_3},\ \text{respectively } \mathbb{P}\,(L_{n+1} = e_3 / L_n = e_1) = \frac{k_3}{k_2 + k_3}.$$

Similar constructions are true for the remaining transitions. Obviously, this does not correspond to the independent framework since the transition matrix

$$\mathbb{Q}^0 = \begin{pmatrix} 0 & \frac{k_2}{k_2+k_3} & \frac{k_3}{k_2+k_3} & 0 \\ \frac{k_{-2}}{k_{-2}+k_4} & 0 & 0 & \frac{k_4}{k_{-2}+k_4} \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{2} & 0 & 0 & \frac{1}{2} \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}.$$

Nevertheless, we shall assume that the host is at equilibrium prior to infection and it is easy to see that the unique invariant measure is the uniform law given by

$$q_1 = q_2 = q_3 = q_4 = \frac{1}{4}. \tag{18}$$

The updating matrices A_n To the transitions $2R_1 \overset{K_1}{\rightleftarrows} R_2$ and $R_1 \overset{k_d}{\to}$ one usually associates the ordinary differential equations

$$\frac{dr_1}{dt} = -2k_1 r_1^2 - k_d r_1 + 2k_{-1} r_2, \ \frac{dr_2}{dt} = k_1 r_1^2 - k_{-1} r_2.$$

By writing down the associated Jacobian matrix at some point $r^0 = \left(r_1^0, r_2^0\right)$, one gets a first-order approximation $\Delta r = \begin{pmatrix} -4k_1 r_1^0 - k_d & 2k_{-1} \\ 2k_1 r_1^0 & -k_{-1} \end{pmatrix} r$. In other words, one constructs the matrix

$$A = I_{2\times 2} + \Delta r = \begin{pmatrix} 1 - 4k_1 r_1^0 - k_d & 2k_{-1} \\ 2k_1 r_1^0 & 1 - k_{-1} \end{pmatrix}.$$

If $r_1^0 = 0$, then the updating of the dimer is done independently of the repressor which is not very realistic. For our toy model, we consider $2k_1 r_1^0 = k_d = k_{-1} = \frac{1}{4}$ i.e.

$$A = \begin{pmatrix} \frac{1}{4} & \frac{1}{2} \\ \frac{1}{4} & \frac{3}{4} \end{pmatrix}. \tag{19}$$

In this framework, 25% of the repressor molecules are degraded (k_d), 50% (i.e $4k_1 r_1^0$) bind together to produce a total of $\frac{4k_1 r_1^0}{2} r_1$ dimers and 25% remain unaltered. For the dimer, 25% (i.e. k_{-1}) break to produce $2k_{-1} r_2$ repressors and 75% remain unaltered.

Remark 24 *Another way of defining A_n would be to keep for Δr the actual Jacobian evaluated at the expectation of uncontrolled X_n i.e. $A_n = \begin{pmatrix} 1 - 4k_1 \mathbb{E}\left[X_n^1\right] - k_d & 2k_{-1} \\ 2k_1 \mathbb{E}\left[X_n^1\right] & 1 - k_{-1} \end{pmatrix}$, then compute $\begin{pmatrix} \mathbb{E}\left[X_n^1\right] \\ \mathbb{E}\left[X_n^2\right] \end{pmatrix} = A_{n-1} \begin{pmatrix} \mathbb{E}\left[X_{n-1}^1\right] \\ \mathbb{E}\left[X_{n-1}^2\right] \end{pmatrix}$, etc.*

The Multiplicative Noise Changes in the trend component have an effect on the couple repressor/dimer in the transcription phase $DR_2 + P \overset{K_t}{\to} DR_2 + P + rR_1$. A careful look at the biochemical reactions shows that binding to the promoter site needs a dimer per binding. Since the DNA mechanism is assumed to be at equilibrium, the number of "averaged" occupied promoter sites can be set proportional to R_2. This reaction will result in a production of r copies per existing dimer as soon as the trend is set to e_2. This implies that

$$C_{2,n} = \begin{pmatrix} 0 & r \\ 0 & 0 \end{pmatrix}. \tag{20}$$

The remaining states induce no noise i.e.

$$C_{i,n} = 0_{2\times 2}, \text{ for } i \in \{1, 3, 4\}. \tag{21}$$

Again, in to simplify the arguments, we assume $r = 1$. We also drop the dependency on n.

We deal with a scaled repressor/dimer component and this is why we add these as pure jump zero-mean contributions. Alternate models are available (see, for example [13]).

Control vs. Controls At this point, we envisage four scenarios concerning the couple repressor/dimer : no external control, control only the dimer, (same one-dimensional) control on both the

dimer and repressor or control (two-dimensional) on each state. To control the dimer, respectively mutually control the couple repressor/dimer, one uses $b_1 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$, respectively $b_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$. To control the two components independently, one uses $[b_1, b_2] \in \mathbb{R}^{2\times 2}$ (which is equivalent, up to renaming the controls, to the use of $B = I_{2\times 2}$). Note that these scenarios correspond to selecting a subset of $\{1, 2\}$.

When $B = b_1$, one computes $\left(A^T\right)^{-1} = \begin{pmatrix} 12 & -4 \\ -8 & 4 \end{pmatrix}$ and $\left(A^T\right)^{-1} C_2^T = \begin{pmatrix} -4 & 0 \\ 4 & 0 \end{pmatrix}$ and notes that $\ker B^T$ is $\left(A^T\right)^{-1} + 2\left(A^T\right)^{-1} C_2^T$-invariant and, thus, $\left(\left(A^T\right)^{-1}; \left(A^T\right)^{-1} \mathcal{C}^T\right)$ −strictly invariant (with the notations of Proposition 17). Therefore, the system is not null-controllable according to Proposition 17.

When $B = b_2$, we compute the explicit solution of (1) associated to a particular choice of the control as follows. For every $x = \begin{pmatrix} x^1 \\ x^2 \end{pmatrix} \in \mathbb{R}^2$, we set

$$u_1 = -\frac{1}{4}x^1 - \frac{3}{4}x^2 \text{ and } u_2 = -\frac{1}{4}\left(\langle L_1, e_2\rangle - \frac{1}{2}\right)x^2,$$

to get

$$X_1^{x,u} = \begin{pmatrix} \left(\langle L_1, e_2\rangle - \frac{1}{2}\right)x^2 \\ 0 \end{pmatrix},\ X_2^{x,u} = 0_{2\times 1}.$$

Then, the system is null-controllable.

Remark 25 *Alternatively, one can use Proposition 10 and compute* $\liminf_{\varepsilon\to 0+} P_0^\varepsilon$ *starting from* $P_2^\varepsilon = 0_{2\times 2}$. *For small values of* $\varepsilon \simeq 10^{-10}$, *numerical values stabilize around* $\begin{pmatrix} 592 & -192 \\ -192 & 64 \end{pmatrix}$ *which is positive-definite. Its minimal eigenvalue is* $\|B(\{2\})\|_{ctrl}^{spec} = \|b_2\|_{ctrl}^{spec} \simeq 1.5647078$.

The system is also controllable if $B = [b_1, b_2]$. In view of Definition 22, it follows that $\mathcal{I} = \{2\}$ provides a minimal efficient intervention plan. In this case, rank and spectral control norms provide the same (unique) answer.

Finally, let us mention that similar procedures can be envisaged for objective-based systems reduction. In this case, the decision is made at subgraph selection level : what reactions to be suppressed and what reactions to be added to preserve a given property. The aim in this framework is to give the smallest set of reactions allowing to achieve the goal. This makes the object of on-going research in both discrete and continuous framework.

4 Proofs of Theorems 2 and 8, Solvability Propositions 10 and 12 and Necessary Condition (Proposition 17)

4.1 Proof of the Main Results

We begin with the proof of the duality-based characterization of controllability concepts.

Proof of Theorem 2. The first two assertions are direct consequences of the duality properties.

One easily notes that

$$\begin{aligned}
&\mathbb{E}\left[\left\langle X_{n+1}^{x,u}, Y_{n+1}^{N,\xi}\right\rangle / \mathcal{F}_n\right] \\
&= \left\langle A_n X_n^{x,u} + Bu_{n+1}, \mathbb{E}\left[Y_{n+1}^{N,\xi}/\mathcal{F}_n\right]\right\rangle + \left\langle X_n^{x,u}, \mathbb{E}\left[\sum_{i=1}^{p} \langle \Delta M_{n+1}, e_i\rangle C_{i,n}^T Z_n^{N,\xi} \Delta M_{n+1}/\mathcal{F}_n\right]\right\rangle \\
&= \left\langle X_n^{x,u}, A_n^T \mathbb{E}\left[Y_{n+1}^{N,\xi}/\mathcal{F}_n\right] + \sum_{i=1}^{p} C_{i,n}^T Z_n^{N,\xi} \mathbb{E}\left[\langle \Delta M_{n+1}, e_i\rangle \Delta M_{n+1}/\mathcal{F}_n\right]\right\rangle + \left\langle Bu_{n+1}, \mathbb{E}\left[Y_{n+1}^{N,\xi}/\mathcal{F}_n\right]\right\rangle \\
&= \left\langle X_n^{x,u}, Y_n^{N,\xi}\right\rangle + \left\langle Bu_{n+1}, \mathbb{E}\left[Y_{n+1}^{N,\xi}/\mathcal{F}_n\right]\right\rangle.
\end{aligned}$$

Hence, by iterating, one gets

$$\mathbb{E}\left[\left\langle X_N^{x,u}, Y_N^{N,\xi}\right\rangle\right] = \left\langle x, Y_0^{N,\xi}\right\rangle + \sum_{n=0}^{N-1} \mathbb{E}\left[\left\langle u_{n+1}, B^T \mathbb{E}\left[Y_{n+1}^{N,\xi}/\mathcal{F}_n\right]\right\rangle\right]. \tag{22}$$

One proceeds in a classical manner by considering the two linear operators

$$\begin{aligned}
&R_N^1 : \mathcal{P}red \longrightarrow \mathbb{L}^2\left(\Omega, \mathcal{F}_N, \mathbb{P}; \mathbb{R}^m\right), \; R_N^1(u) = X_N^{0,u}, \text{ for all } u \in \mathcal{P}red, \\
&R_N^2 : \mathbb{R}^m \longrightarrow \mathbb{L}^2\left(\Omega, \mathcal{F}_N, \mathbb{P}; \mathbb{R}^m\right), \; R_N^2(x) = X_N^{x,0}, \text{ for all } x \in \mathbb{R}^m.
\end{aligned}$$

The reader is invited to recall that $\mathcal{P}red$ stands for the family of $\mathbb{R}^d$-valued, $\mathbb{F}$-predictable controls. (It is considered here as a subspace of product of $\mathbb{L}^2\left(\Omega, \mathcal{F}_n, \mathbb{P}; \mathbb{R}^d\right)$-spaces.) In view of (22), the adjoints of the linear operators are given by

$$\left(R_N^1\right)^*(\xi) = \left(B^T \mathbb{E}\left[Y_n^{N,\xi}/\mathcal{F}_{n-1}\right]\right)_{n\geq 1}, \;\; \left(R_N^2\right)^*(\xi) = Y_0^{N,\xi}, \tag{23}$$

for all $\xi \in \mathbb{L}^2\left(\Omega, \mathcal{F}_N, \mathbb{P}; \mathbb{R}^m\right)$. Then, approximate null-controllability is equivalent to the image (range) inclusion $\mathrm{Im}\left(R_N^2\right) \subset cl\left(\mathrm{Im}\left(R_N^1\right)\right)$, where cl is the usual Kuratowski closure operator. Furthermore, this is equivalent to the kernel inclusion $\ker\left(\left(R_N^1\right)^*\right) \subset \ker\left(\left(R_N^2\right)^*\right)$ which leads to the second assertion. Similar, approximate controllability is equivalent to $cl\left(\mathrm{Im}\left(R_N^1\right)\right) = \mathbb{L}^2\left(\Omega, \mathcal{F}_N, \mathbb{P}; \mathbb{R}^m\right)$. Hence, it is equivalent to $\ker\left(\left(R_N^1\right)^*\right) = \{0\}$ which leads to the first assertion.

For the third assertion, since Ω is assumed to be the sample space, it follows that $\mathbb{L}^2\left(\Omega, \mathcal{F}_N, \mathbb{P}; \mathbb{R}^m\right)$ can be seen as $\mathbb{R}^{mp^N}$. Hence, the linear subspace $\mathrm{Im}\left(R_N^1\right)$ is finite-dimensional and, thus, $cl\left(\mathrm{Im}\left(R_N^1\right)\right) = \mathrm{Im}\left(R_N^1\right)$. In this case, approximate null-controllability is written down as $\mathrm{Im}\left(R_N^2\right) \subset \mathrm{Im}\left(R_N^1\right)$ (i.e. one actually has exact null-controllability), or, equivalently (see, for example, [26, Appendix B, Proposition B.1]),

$$\left|\left(R_N^2\right)^* \xi\right| \leq k \left\|\left(R_N^1\right)^* \xi\right\|_{\mathcal{P}red}, \text{ for some } k > 0.$$

The necessary and sufficient condition (3) follows from (23). ■

We now give the proof of the second main result of the paper providing the link between the controllability (pseudo)norm and the backward stochastic Riccati difference scheme.

Proof of Theorem 8. For the first assertion, using the particular form of $\alpha_{n,\varepsilon}$ and $\eta_{n,\varepsilon}$, one simply writes down

$$\begin{aligned}
&\langle P_n^\varepsilon y_n^{y_0,v}, y_n^{y_0,v}\rangle \\
&= \mathbb{E}\left[\left\langle P_{n+1}^\varepsilon \left[A_n^T\right]^{-1} y_n^{y_0,v}, \left[A_n^T\right]^{-1} y_n^{y_0,v}\right\rangle / \mathcal{F}_n\right] + \left|B^T \left[A_n^T\right]^{-1} y_n^{y_0,v}\right|^2 - \left\langle \alpha_{n,\varepsilon}^T \eta_{n,\varepsilon}^{-1} \alpha_{n,\varepsilon} y_n^{y_0,v}, y_n^{y_0,v}\right\rangle \\
&= \mathbb{E}\left[\left\langle P_{n+1}^\varepsilon y_{n+1}^{y_0,v}, y_{n+1}^{y_0,v}\right\rangle / \mathcal{F}_n\right] + \varepsilon |v_{n+1}|^2 + \left|B^T \mathbb{E}\left[y_{n+1}^{y_0,v}/\mathcal{F}_n\right]\right|^2 - \left|\eta_{n,\varepsilon}^{-1/2} \alpha_{n,\varepsilon} y_n^{y_0,v} - \eta_{n,\varepsilon}^{1/2} v_{n+1}\right|^2.
\end{aligned}$$

By iterating and taking expectation, one gets

$$\langle P_0^\varepsilon y_0, y_0\rangle = \varepsilon \sum_{n=0}^{N-1} \mathbb{E}\left[|v_{n+1}|^2\right] + \mathbb{E}\left[\sum_{n=0}^{N-1} \left|B^T \mathbb{E}\left[y_{n+1}^{y_0,v}/\mathcal{F}_n\right]\right|^2\right]$$

$$- \sum_{n=0}^{N-1} \mathbb{E}\left[\left|\eta_{n,\varepsilon}^{-1/2}\delta_{n,\varepsilon} y_n^{y_0,v} - \eta_{n,\varepsilon}^{1/2} v_{n+1}\right|^2\right]. \tag{24}$$

If the system (1) is (approximately) null-controllable, then there exists some positive constant $c > 0$ such that

$$\inf_{(v_n)_{1\leq n\leq N} \mathbb{F}\text{-predictable}} \mathbb{E}\left[\sum_{n=0}^{N-1} \left|B^T \mathbb{E}\left[y_{n+1}^{y_0,v}/\mathcal{F}_n\right]\right|^2\right] \geq c\,|y_0|^2.$$

In particular, by taking the feedback control $v_{n+1}^\varepsilon := \eta_{n,\varepsilon}^{-1}\delta_{n,\varepsilon} y_n^{y_0,v^\varepsilon}$, one establishes that

$$\langle P_0^\varepsilon y_0, y_0\rangle \geq c\,|y_0|^2$$

and the conclusion follows.

Conversely, if $\liminf_{\varepsilon\to 0+} P_0^\varepsilon \geq cI$, for some $c > 0$, then, for every $\varepsilon > 0$ small enough and every predictable control v, one gets

$$\varepsilon \sum_{n=0}^{N-1} \mathbb{E}\left[|v_{n+1}|^2\right] + \mathbb{E}\left[\sum_{n=0}^{N-1} \left|B^T \mathbb{E}\left[y_{n+1}^{y_0,v}/\mathcal{F}_n\right]\right|^2\right] \geq \frac{c}{2}\,|y_0|^2$$

and the conclusion follows by letting $\varepsilon \to 0$.

For the second assertion, one notes that (24) implies that

$$\langle P_0^\varepsilon y_0, y_0\rangle = \inf_{(v_n)_{1\leq n\leq N} \mathbb{F}\text{-predictable}} \left(\varepsilon \sum_{n=0}^{N-1} \mathbb{E}\left[|v_{n+1}|^2\right] + \mathbb{E}\left[\sum_{n=0}^{N-1} \left|B^T \mathbb{E}\left[y_{n+1}^{y_0,v}/\mathcal{F}_n\right]\right|^2\right]\right)$$

and the conclusion follows by letting $\varepsilon \to 0$. ■

4.2 Proof of the Solvability Results

We begin with the proof for the solvability of the BSRDS in the case of non-random coefficients.

Proof of Proposition 10. The proof is given by (descending) induction. For $n = N$, it is clear that $Q_{N-1}^\varepsilon = 0$. Since A_{N-1} and $\mathcal{C}_{N-1}$ are deterministic, it is clear that the iterative step defining P_{N-1}^ε in scheme (8) reduces to (9). Let us assume that P_{n+1}^ε has been constructed according to this deterministic scheme and is a positive-semidefinite (non-random)matrix. Then, $Q_n^\varepsilon = 0$. Since A_n and C_n are deterministic, the definition of P_n^ε according to scheme (8) reduces to (9). We only need to prove that this latter scheme is consistent and provides positive-semidefinite matrices. One begin with noting that as soon as P_{n+1}^ε is positive-semidefinite, the matrix

$$\Delta_{n+1}^\varepsilon := \left(q_j\left(\delta_{j,k} - q_k\right) P_{n+1}^\varepsilon\right)_{1\leq j,k\leq p} \in \mathbb{R}^{mp\times mp}$$

is also positive-semidefinite. Indeed, it suffices to set $DD^T = \left(q_j\left(\delta_{j,k} - q_k\right)\right)_{1\leq j,k\leq p}$ given by Cholesky decomposition and $\mathcal{D}_{j,k} := D_{j,k} I_{m\times m}$, for all $1 \leq j, k \leq p$. Then

$$\Delta_{n+1}^\varepsilon = \mathcal{D}\begin{pmatrix} P_{n+1}^\varepsilon & 0_{m\times m} & 0_{m\times m} & \dots & 0_{m\times m} \\ 0_{m\times m} & P_{n+1}^\varepsilon & 0_{m\times m} & \dots & 0_{m\times m} \\ 0_{m\times m} & 0_{m\times m} & P_{n+1}^\varepsilon & \dots & 0_{m\times m} \\ \dots & \dots & \dots & \dots & \dots \\ 0_{m\times m} & 0_{m\times m} & 0_{m\times m} & \dots & P_{n+1}^\varepsilon \end{pmatrix}\mathcal{D}^T$$

is obviously positive-semidefinite. It follows that $\eta_{n,\varepsilon}$ given by (9) is positive-definite. Second, using a classical intuition on feedback optimal control, one writes

$$\begin{aligned}
&A_n^{-1}\left(P_{n+1}^{\varepsilon}+BB^T\right)\left[A_n^T\right]^{-1}-\alpha_{n,\varepsilon}^T\eta_{n,\varepsilon}^{-1}\alpha_{n,\varepsilon}\\
&=\left[A_n^{-1}-\alpha_{n,\varepsilon}^T\eta_{n,\varepsilon}^{-1}\mathcal{C}_nA_n^{-1}\right]\left(P_{n+1}^{\varepsilon}+BB^T\right)\left[\left[A_n^{-1}\right]^T-\left[A_n^{-1}\right]^T\mathcal{C}_n\eta_{n,\varepsilon}^{-1}\alpha_{n,\varepsilon}\right]\\
&+\alpha_{n,\varepsilon}^T\eta_{n,\varepsilon}^{-1}\left(\varepsilon I_{mp\times mp}+\Delta_{n+1}^{\varepsilon}\right)\eta_{n,\varepsilon}^{-1}\alpha_{n,\varepsilon}.
\end{aligned}$$

This implies that P_n^{ε} is positive-semidefinite whenever P_{n+1}^{ε} is positive-semidefinite and the induction step is complete. ■

The second proof concerns the solvability of the BSRDS in the absence of multiplicative noise (i.e. $C=0$).

Proof of Proposition 12. One begins with setting

$$p_N^{\varepsilon}=0 \text{ and } q_N^{\varepsilon}=0$$

and notes that P_{N-1}^{ε} given by (10) satisfies

$$P_{N-1}^{\varepsilon}=A_{N-1}^{-1}BB^T\left[A_{N-1}^T\right]^{-1}=A^{-1}\left(N-1,L_{N-1}\right)\left(p_N^{\varepsilon}+BB^T-q_N^{\varepsilon}\right)\left(A^{-1}\left(N-1,L_{N-1}\right)\right)^T.$$

Next, one recalls that $Q_{N-2}=\left[Q_{N-2,1}\ Q_{N-2,2}\ ...\ Q_{N-2,m}\right]$, where $Q_{N-2,j}\in\mathbb{R}^{m\times p}$. One easily computes

$$\left[Q_{N-2,1}^{\cdot,l},Q_{N-2,2}^{\cdot,l},...,Q_{N-2,m}^{\cdot,l}\right]=A^{-1}\left(N-1,e_l\right)\left(p_N^{\varepsilon}+BB^T-q_N^{\varepsilon}\right)\left(A^{-1}\left(N-1,e_l\right)\right)^T,$$

for all $1\leq l\leq p$. Therefore, the conclusion holds true for $n=N-1$. We proceed by (decreasing) induction and assume the conclusion to hold true for $n+1$ and prove it for $n\leq N-2$. One easily notes that,due to the recurrence assumption, the following equalities hold true for α and η computed as in (10).

$$\begin{cases}
\alpha_{n,\varepsilon}^j=-\overline{\alpha}_{n,\varepsilon}^j\left[A_n^T\right]^{-1},\ \text{where}\\
\overline{\alpha}_{n,\varepsilon}^j=\left[\sum_{l=1}^{p}q_l\left(\delta_{j,l}-q_j\right)A^{-1}\left(n+1,e_l\right)\left(p_{n+2}^{\varepsilon}+BB^T-q_{n+2}^{\varepsilon}\right)\left(A^{-1}\left(n+1,e_l\right)\right)^T\right]\ \text{and}\\
\eta_{n,\varepsilon}^{j,k}=\varepsilon\delta_{j,k}I_{m\times m}+\sum_{l=1}^{p}q_l\left(q_j-\delta_{j,l}\right)\left(q_k-\delta_{k,l}\right)A^{-1}\left(n+1,e_l\right)\left(p_{n+2}^{\varepsilon}+BB^T-q_{n+2}^{\varepsilon}\right)\left(A^{-1}\left(n+1,e_l\right)\right)^T,
\end{cases}$$

for all $1\leq j,k\leq p$. We set

$$(25)\quad\begin{cases}
p_{n+1}^{\varepsilon}:=\mathbb{E}\left[P_{n+1}^{\varepsilon}/\mathcal{F}_n\right]=\sum_{l=1}^{p}q_lA^{-1}\left(n+1,e_l\right)\left(p_{n+2}^{\varepsilon}+BB^T-q_{n+2}^{\varepsilon}\right)\left(A^{-1}\left(n+1,e_l\right)\right)^T,\\
q_{n+1}^{\varepsilon}=\overline{\alpha}_{n,\varepsilon}^T\eta_{n,\varepsilon}^{-1}\overline{\alpha}_{n,\varepsilon}.
\end{cases}$$

We will see in just one moment that $\eta_{n,\varepsilon}^{-1}$ (hence, q_{n+1}^{ε}) is consistent. By (10), it follows that

$$P_n^{\varepsilon}=A_n^{-1}\left(p_{n+1}^{\varepsilon}+BB^T-q_{n+1}^{\varepsilon}\right)\left[A_n^T\right]^{-1}.$$

For Q, the assertion is obtained as in the first step. We come back to the quantities p_{n+1}^{ε} and q_{n+1}^{ε} given by (25) and show that they are well-defined and satisfy (11). The induction assumption $p_{n+2}^{\varepsilon}\geq q_{n+2}^{\varepsilon}$ implies that p_{n+1}^{ε} is positive-semidifinite. Second, with the notations

$$\begin{aligned}
\mathcal{A}&:=\left(\sqrt{q_1}A^{-1}\left(n+1,e_1\right),...,\sqrt{q_p}A^{-1}\left(n+1,e_p\right)\right)\in\mathbb{R}^{m\times mp},\\
\mathcal{P}&:=\left(\delta_{j,k}\left(p_{n+2}^{\varepsilon}+BB^T-q_{n+2}^{\varepsilon}\right)\right)_{1\leq j,k\leq p}\in\mathbb{R}^{mp\times mp},\\
\mathcal{D}&:=\left(\sqrt{q_k}\left(\delta_{j,k}-q_j\right)A^{-1}\left(n+1,e_k\right)\right)_{1\leq j,k\leq p}\in\mathbb{R}^{mp\times mp},
\end{aligned}$$

one has

$$\eta_{n,\varepsilon} = \varepsilon I_{mp\times mp} + \mathcal{D}\mathcal{P}\mathcal{D}^T > 0 \text{ and } \overline{\alpha}_{n,\varepsilon} = \mathcal{D}\mathcal{P}\mathcal{A}^T.$$

For the inequality, we have used the induction assumption $p^{\varepsilon}_{n+2} \geq q^{\varepsilon}_{n+2}$. As consequence, q^{ε}_{n+1} is well-defined and positive-semidefinite. Finally,

$$\begin{aligned}
&p^{\varepsilon}_{n+1} - q^{\varepsilon}_{n+1} = \mathcal{A}\mathcal{P}\mathcal{A}^T - \mathcal{A}\mathcal{P}\mathcal{D}^T\left(\varepsilon I_{mp\times mp} + \mathcal{D}\mathcal{P}\mathcal{D}^T\right)^{-1}\mathcal{D}\mathcal{P}\mathcal{A}^T \\
&= \left(\mathcal{A} - \mathcal{A}\mathcal{P}\mathcal{D}^T\left(\varepsilon I_{mp\times mp} + \mathcal{D}\mathcal{P}\mathcal{D}^T\right)^{-1}\mathcal{D}\right)\mathcal{P}\left(\mathcal{A} - \mathcal{A}\mathcal{P}\mathcal{D}^T\left(\varepsilon I_{mp\times mp} + \mathcal{D}\mathcal{P}\mathcal{D}^T\right)^{-1}\mathcal{D}\right)^T \\
&+ \varepsilon\mathcal{A}\mathcal{P}\mathcal{D}^T\left(\varepsilon I_{mp\times mp} + \mathcal{D}\mathcal{P}\mathcal{D}^T\right)^{-1}\left(\varepsilon I_{mp\times mp} + \mathcal{D}\mathcal{P}\mathcal{D}^T\right)^{-1}\mathcal{D}\mathcal{P}\mathcal{A}^T,
\end{aligned}$$

which is clearly positive-semidefinite. Our induction step is now complete and the conclusion follows. ■

4.3 Proof of Proposition 17

Proof of Proposition 17. Let us denote by $\mathcal{V}$ the largest subspace of $\ker\left(B^T\right)$ which is $\left(\left[A^T\right]^{-1}; \left(\mathcal{C}(1)A^{-1}\right)^T, ..., \left(\mathcal{C}(p)A^{-1}\right)^T\right)$ −strictly invariant. According to [28, Theorem 3.2] (see also [8, Lemma 4.6]), the set $\mathcal{V}$ is equally $\left(\left[A^T\right]^{-1}; \left(\mathcal{C}(1)A^{-1}\right)^T\Pi_{\mathcal{V}}, ..., \left(\mathcal{C}(p)A^{-1}\right)^T\Pi_{\mathcal{V}}\right)$ −feedback invariant. Thus, there exists a family of linear operators $(F(l))_{1\leq l\leq p} \subset \mathbb{R}^{m\times m}$ such that $\mathcal{V}$ is $\left(\left[A^T\right]^{-1} + \sum_{l=1}^{p}\left(\mathcal{C}(l)A^{-1}\right)^T\Pi_{\mathcal{V}}F(l)\right)$- invariant. We consider the linear stochastic system

$$\begin{cases} x^{y_0}_{n+1} := \left(\left[A^T\right]^{-1} + \sum_{l=1}^{p}\left(\mathcal{C}(l)A^{-1}\right)^T\Pi_{\mathcal{V}}F(l)\right)x^{y_0}_n + \sum\limits_{l=1}^{p}\langle\Delta M_{n+1}, e_l\rangle\,\Pi_{\mathcal{V}}F(l)\,x^{y_0}_n, \\ x^{y_0}_0 = y_0. \end{cases}$$

Then $x^{y_0}_{n+1}$ coincides with the solution of (4) associated to the feedback control $v^{feedback}(n+1, y) = [\Pi_{\mathcal{V}}F(1)y, ..., \Pi_{\mathcal{V}}F(p)y]$, for all $n \geq 0$. Moreover, whenever $y_0 \in \mathcal{V}$, one gets $x^{y_0}_{n+1} \in \mathcal{V}$, $\mathbb{P}-a.s.$ for all $n \geq 0$. In particular, recalling that $\mathcal{V} \subset \ker\left(B^T\right)$, it follows that $B^T\mathbb{E}\left[y^{y_0,v^{feedback}}_{n+1}/\mathcal{F}_n\right] = 0$, $\mathbb{P}-a.s.$ for all $n \geq 0$. By our controllability assumption and Criterion 4, one deduces $y_0 = 0$ and our Proposition is complete by recalling that $y_0 \in \mathcal{V}$ is arbitrary. ■

References

[1] V. Barbu, A. Răşcanu, and G. Tessitore. Carleman estimates and controllability of linear stochastic heat equations. *Appl. Math. Optim.*, 47(2):97–120, 2003.

[2] R. Buckdahn, M. Quincampoix, and G. Tessitore. A characterization of approximately controllable linear stochastic differential equations. In *Stochastic partial differential equations and applications—VII*, volume 245 of *Lect. Notes Pure Appl. Math.*, pages 53–60. Chapman & Hall/CRC, Boca Raton, FL, 2006.

[3] Samuel N. Cohen and Robert J. Elliott. A general theory of finite state backward stochastic difference equations. *Stochastic Processes and their Applications*, 120(4):442 – 466, 2010.

[4] Samuel N. Cohen and Robert J. Elliott. Backward stochastic difference equations and nearly time-consistent nonlinear expectations. *SIAM Journal on Control and Optimization*, 49(1):125–139, 2011.

[5] SamuelN. Cohen and RobertJ. Elliott. Backward stochastic difference equations with finite states. In Arturo Kohatsu-Higa, Nicolas Privault, and Shuenn-Jyi Sheu, editors, *Stochastic Analysis with Financial Applications*, volume 65 of *Progress in Probability*, pages 33–42. Springer Basel, 2011.

[6] O.L.V. Costa, M.D. Fragoso, and R.P. Marques. *Discrete-Time Markov Jump Linear Systems*. Probability and Its Applications. Springer-Verlag London, 2005.

[7] Oswaldo L.V. Costa and Alexandre de Oliveira. Optimal mean–variance control for discrete-time linear systems with markovian jumps and multiplicative noises. *Automatica*, 48(2):304 – 315, 2012.

[8] R. F. Curtain. Invariance concepts in infinite dimensions. *SIAM J. Control and Optim.*, 24(5):1009–1030, SEP 1986.

[9] E. Fernández-Cara, M. J. Garrido-Atienza, and J. Real. On the approximate controllability of a stochastic parabolic equation with a multiplicative noise. *C. R. Acad. Sci. Paris Sér. I Math.*, 328(8):675–680, 1999.

[10] Daniel T. Gillespie. Exact stochastic simulation of coupled chemical reactions. *The Journal of Physical Chemistry*, 81(25):2340–2361, 1977.

[11] D. Goreac. Approximate controllability for linear stochastic differential equations in infinite dimensions. *Applied Mathematics and Optimization*, 60(1):105–132, 2009.

[12] D. Goreac. Controllability properties of linear mean-field stochastic systems. *Stochastic Analysis and Applications*, 32(02):280–297, 2014.

[13] D. Goreac. Asymptotic control for a class of piecewise deterministic markov processes associated to temperate viruses. *SIAM Journal on Control and Optimization*, 53(4):1860–1891, 2015.

[14] Dan Goreac. A kalman-type condition for stochastic approximate controllability. *Comptes Rendus Mathematique*, 346(3–4):183 – 188, 2008.

[15] Dan Goreac. A note on the controllability of jump diffusions with linear coefficients. *IMA Journal of Mathematical Control and Information*, 29(3):427–435, 2012.

[16] Dan Goreac and Miguel Martinez. Algebraic invariance conditions in the study of approximate (null-)controllability of markov switch processes. *Mathematics of Control, Signals, and Systems*, pages 1–28, 2015, 10.1007/s00498-015-0146-1.

[17] J. Hasty, J. Pradines, M. Dolnik, and J.J. Collins. Noise-based switches and amplifiers for gene expression. *PNAS*, 97(5):2075–2080, 2000.

[18] M. L. J. Hautus. Controllability and observability conditions of linear autonomous systems. *Nederl. Akad. Wetensch. Proc. Ser. A 72 Indag. Math.*, 31:443–448, 1969.

[19] Birgit Jacob and Jonathan R. Partington. On controllability of diagonal systems with one-dimensional input space. *Systems & Control Letters*, 55(4):321 – 328, 2006.

[20] Birgit Jacob and Hans Zwart. Exact observability of diagonal systems with a finite-dimensional output operator. *Systems & Control Letters*, 43(2):101 – 109, 2001.

[21] L. Lovász. Submodular functions and convexity. In Achim Bachem, Bernhard Korte, and Martin Grötschel, editors, *Mathematical Programming The State of the Art*, pages 235–257. Springer Berlin Heidelberg, 1983.

[22] R. C. Merton. Option pricing when underlying stock returns are discontinuous. *J. Financ. Econ.*, 3:125– 144, 1976.

[23] G.L. Nemhauser, L.A. Wolsey, and M.L. Fisher. An analysis of approximations for maximizing submodular set functions—i. *Mathematical Programming*, 14(1):265–294, 1978.

[24] E. Pardoux and S.G. Peng. Adapted solution of a backward stochastic differential equation. *Syst. Control Lett.*, 14(1):55 – 61, 1990.

[25] S. Peng. Backward stochastic differential equation and exact controllability of stochastic control systems. *Progr. Natur. Sci. (English Ed.)*, 4:274–284, 1994.

[26] G. Da Prato and J. Zabczyk. *Stochastic Equations in Infinite Dimensions*. Cambridge University Press, 1992. Cambridge Books Online.

[27] D. Russell and G. Weiss. A general necessary condition for exact observability. *SIAM Journal on Control and Optimization*, 32(1):1–23, 1994.

[28] E. J. P. G. Schmidt and R. J. Stern. Invariance theory for infinite dimensional linear control systems. *Applied Mathematics and Optimization*, 6(2):113–122, 1980.

[29] M. Sirbu and G. Tessitore. Null controllability of an infinite dimensional SDE with state- and control-dependent noise. *Systems & Control Letters*, 44(5):385–394, DEC 14 2001.

[30] T. Summers, F. Cortesi, and J. Lygeros. On submodularity and controllability in complex dynamical networks. *IEEE Transactions on Control of Network Systems*, 2014.

[31] J. Yong and X.Y. Zhou. *Stochastic Controls. Hamiltonian Systems and HJB Equations*. Springer-Verlag, New York, 1999.